(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number
WO 01/80540 A1

(51) International Patent Classification[7]: H04M 3/00

(21) International Application Number: PCT/US01/07457

(22) International Filing Date: 7 March 2001 (07.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/550,006 14 April 2000 (14.04.2000) US

(71) Applicant: GENESYS TELECOMMUNICATION LABORATORIES, INC. [US/US]; 11th floor, 1155 Market Street, San Francisco, CA 94103 (US).

(72) Inventors: ANISIMOV, Nikolay; Apartment #64, 1450 Creekside Drive, Walnut Creek, CA 94596 (US). KISHINSKY, Konstatin; 10 Whitman Court, San Carlos, CA 94070 (US). KOROLEV, Nikolay; Apartment #2028, 1469 Creekside Drive, Walnut Creek, CA 94596 (US). POGOSSIANTS, Gregory; 669 Maybell Avenue, Palo Alto, CA 94306 (US). SHTIVELMAN, Yuri; 801 Marina Boulevard, San Francisco, CA 94123 (US). Zyarko, Vadim; Apartment #42, 1440 Creekside Drive, walnut Creek, CA 94596 (US).

(74) Agent: BOYS, Donald, R.; P.O. Box 187, Aromas, CA 95004 (US).

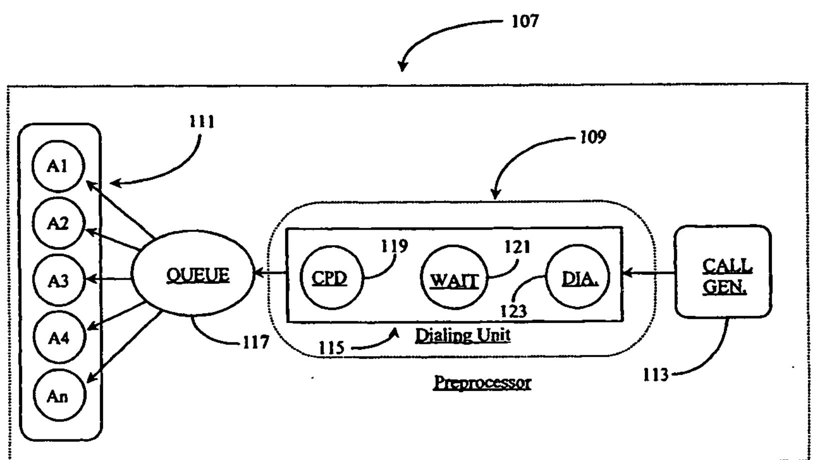
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

[Continued on next page]

(54) **Title:** METHOD FOR PREDICTING AND MANAGING CALL LOAD BY DETERMINING THE OPTIMUM FREQUENCY OF OUTBOUND CALL GENERATION



Communication Center (Out-Bound Campaign)

(57) **Abstract:** A system for balancing outbound dialing rate with agent utilization in a telephony call center, minimizing wait time for answered outdialed calls has a call number generating module (113) for generating numbers to be automatically dialed, a dialing unit (123) for dialing numbers generated, a queue (117) for queuing answered calls, and a stat module for monitoring performance and generating a call generation rate. The system uses an analytical method or a simulation method for determining the call generation rate. In the analytical method, one or more of distribution function of system processes, numerical methods, solving of non-linear equations, or probability techniques is used. In the simulation method, parameters relating to a queing system are estimated in absence of sufficient information by simulating each of the system's processes, predicting behavior of the system on basis of previous experience, and finding optimal point in future to make a next call.

WO 01/80540 A1

Method for Predicting and Managing Call Load by Determining the Optimum Frequency of Outbound Call Generation.

Field of the Invention

The present invention is in the field of Computer Telephony Integrated (CTI) communication systems including both connection-oriented, switched telephony (COST) systems and Data Network Telephony (DNT) systems such as Internet-Protocol-Network-Telephony (IPNT) systems, and pertains more particularly to methods and apparatus for predicting an optimum frequency for an out-bound call generator in skill-based agent level routing (ALR) environments.

Cross-Reference to Related Documents

The present invention is a continuation in part of U.S. patent S/N 09/209,306 entitled ***"Method for Estimating Telephony System-Queue Waiting Time in an Agent Level Routing Environment"*** filed on 12/11/98 disclosure of which is incorporated herein by reference.

Background of the Invention

Telephone call processing and switching systems are, at the time of the present patent application, relatively sophisticated, computerized systems, and development and introduction of new systems continues.

routing to an agent will be at the call center. If there are several call centers, the organization may provide several numbers, one for each call center, and the customer may be expected to use the number for the closest center, or for the center advertised to provide specifically the service he or she might need. In many cases the number provided will connect the caller with a first Service Control Point (SCP) which is adapted to pre-process incoming calls and forward the calls to call centers.

Routing of calls, then, may be on several levels. Pre-routing may be done at SCPs and further routing may be, and often is, accomplished at individual call centers. As described above, a call center typically involves a central switch, typically including an Automatic Call Distributor (ACD). The central switch is connected to the PSTN or other call network, as is well-known in the art. Agents, trained to interact with callers, service telephones connected to the central switch.

If the call center consists of just a central switch and connected telephone stations, the routing that can be done is quite limited. Current art telephony switches, although increasingly computerized, are limited in the range of computer processes that may be performed. For this reason additional computer capability in the art has been added for such central switches by connecting computer processors, adapted to run control routines and to access databases, to the central switch. The processes of incorporating computer enhancement to telephone switches is known in the art as Computer Telephony Integration (CTI), and the hardware and software together is referred to as CTI equipment. Typically the CTI processor, executing CTI applications, monitors the activity of the switch and status of calls and equipment, and issues instructions and commands to the switch.

There are techniques practiced in the industry aimed at alleviating long caller-queue waiting periods. One such standard development involves call load-balancing wherein incoming calls are distributed (routed) more evenly to available call centers such that queue lengths individual to separate call centers are somewhat the same. This technique may also be practiced in single call centers wherein calls are distributed among separate groups of agents. While this technique helps to even out call loads among different queues, queue length may still be high during peak traffic periods.

Another technique involves transferring a call to an alternate destination when that call approaches a pre-set maximum queue-waiting time for an agent. The alternate destination may help to keep the caller on the line via interactive method such as reviewing the purpose of the call or perhaps advertising products, while waiting for an available agent. However, a long queue can still be an irritating factor for many callers, even when some form of entertainment such as music is provided.

The above-described techniques may help to stabilize overall queue waiting times within call centers, or help to alleviate caller stress when waiting time is excessive, but they only partially address the problem. At peak call-in periods queue waits may still be high even though calls are distributed evenly. Regardless of the distribution (routing) method used, callers are generally not informed of expected waiting time.

With advances in call routing becoming more prevalent in the art advanced techniques must be developed for estimating queue wait times. New technologies include priority queuing, virtual queuing, routing to agents based on skill-set of the agent (e.g. language, level of expertise, etc.), routing to agents based on level or state of availability, routing to agents based on pre-acquired and/or pre-stored caller information, routing to agents based on priority assignment of call, and so on.

In this scenario, there are further problems with which to contend in order to effect adequate call-load to agent utilization ratios. For example, there are missed calls due to non-answer. There will invariably be customer pick-ups that simply hang-up immediately. There will also be calls answered by an answering machine. Therefore, out of a total number of out-bound call attempts, only a certain percentage will translate into calls-in-queue.

The calls-in-queue resulting from an out-bound campaign must be treated differently than normal incoming calls because of the fact that a customer that has been called by the center is much less likely to accept even a small wait time in queue. Therefore, it is desired that there be zero or near zero wait time for connecting out-bound contacts to waiting agents.

The challenge then, is how to manage a call-load resulting from an out-bound campaign such that customers are not kept waiting for any long period, and that agents are still not underutilized in answering calls.

What is clearly needed is a method for predicting call generation rates in an out-bound-call campaign such that the appropriate call-loads are produced for the appropriate number of agents working the calls with customers experiencing very little or no wait time. A system such as this would allow a higher success rate in servicing customers contacted through out-bound call campaigns by enabling agents to retain those contacts normally lost because of hang-up during queue wait time.

Summary of the Invention

In a preferred embodiment of the invention a system for balancing outbound dialing rate with agent utilization in a telephony call center,

In still another aspect a simulation method for minimizing time in queue for parties answering outdialed calls, in an outdialing system is provided comprising steps of (a) simulating each of the system's processes; (b) predicting behavior of the system on basis of previous experience; and (c) finding an optimal point in the future to make a next call, based on results of steps (a) and (b).

In embodiments of the system taught in enabling detail below an outdialing system is provided that performs better than prior art systems, minimizing time in queue for answered calls.

Brief Description of the Drawing Figures

Fig. 1 is a block diagram illustrating a communication system having call-waiting estimation capability according to prior art.

Fig. 2 is an overview of a CTI-enhanced telecommunications system wherein estimated-waiting time (EWT) may be practiced according to an embodiment of the present invention.

Fig. 3 is a table illustrating practice of the present invention in a skill-based priority queue.

Fig. 4 is a block diagram illustrating out-bound call processing system as known to the inventor.

Fig. 5 is a block diagram illustrating variables considered in determining call-generation frequency according to an embodiment of the present invention.

Fig. 6 is a block diagram of an outbound call-processing model according to an embodiment of the present invention.

adapted as a central call-in switch and is a first destination for all calls 16 destined to call-in center 19.

As is known in the art, ACD 21 employs a FIFO queuing system 22 wherein calls 16 wait until an available representative is available to handle the call on a first-in first-out basis. ACD 21 is connected via an internal wiring system 41 to a plurality of telephones 33, 35, 37, and 39 which are implemented one-telephone-per to agent workstations 25, 27, 29, and 31. Call-in center 19 is not CTI enhanced to the extent that agents at workstations 25-31 have access to LAN connected PC's nor is call-in center 19 capable of much intelligent routing such as is possible in CTI enhanced environments. It is to this simple prior art system that EWT is implemented in rather limited scope as described above.

In order to achieve EWT in this prior art system, a call processor 23 is provided as a dedicated unit for estimating waiting time associated with FIFO queue 22. Processor 23 is connected to ACD 21 via a data control line 26. In prior art specification 5,020,095 which was mentioned above with reference to the background section, incoming trunks 17 are diverted through such a processor as processor 23 lending to the dedicated nature of the device as disclosed therein. However, it will be apparent to one with skill in the art that the same level of control over ACD 21 may be provided via control line 26 with the appropriate trunk interfaces installed in ACD 21. Processor 23 would not be considered a CTI processor in current art as intelligent routing applications are not incorporated therein.

Several EWT software routines are provided and installed in processor 23 and adapted, among other purposes, for monitoring and interfacing with calls 16 as they arrive in queue 22. Other capabilities include agent monitoring for busy or not busy, voice interface capability for informing callers of EWT, a means for calculating average call time per call,

prior calls may include one or more abandoned calls thereby producing an unnaturally low average call-handling time that is communicated to the next caller arriving in queue. Moreover, EWT as communicated to a caller must take into account the number of calls ahead of his or her call in queue or N (total calls in queue). If there are many calls ahead multiplied by an unnaturally low average call handling time, the caller may receive a misleading time estimate.

Another problem with prior art as exemplified herein is that the actual time for handling calls may vary widely from call to call. Therefore, taking an average handling time over just a few calls (three in U.S. Patent 5,020,095) is not reflective of a confident average as it is well known that accuracy of any average taken improves with the number of units (in this case calls) to be averaged. Still another state that is not considered in the prior art is the fact that agents in many cases may receive calls from more than one queue. Therefore, simply summing up the number of agents (m) working from one queue will not suffice as a portion of their time may be devoted to answering calls from another queue. Therefore, a more flexible treatment of EWT must be accomplished by way of revised formulas and added software in order to successfully and more accurately practice EWT. Such a flexible implementation of EWT is described in enabling detail below.

Fig. 2 is an overview of a CTI-enhanced telecommunications system 45 wherein EWT may be practiced according to an embodiment of the present invention. System 45 in this example comprises a PSTN network 47, an Internet network 49, and a communication center 51. PSTN 47 may be a public or private COST network as is known in the art. Internet 49 may be of the form of another data-packet network as is known in the art such as a private WAN or corporate Intranet. Communication center 51, in this embodiment, is capable of receiving incoming calls from both PSTN 47 and

the network level via digital network 63, processor 57, and CTI connection 59. This embodiment represents state-of-the-art communication technology on the COST side of communication system 45 as is known and available to the inventor.

Internet 49 is meant to illustrate, for purposes of discussion, that the method and apparatus of the present invention may be adapted and equated to data-network-telephony (DNT), and more particularly Internet-protocol-network-telephony (IPNT) as would be practiced with regard to Internet 49 and communication center 51 in this embodiment. However, for the purpose of this specification, most reference will be to practice of the present invention in a COST network.

Within communication center 51, there is illustrated a plurality of agent stations, station 79 and station 81. Stations 79 and 81 are each adapted and equipped to facilitate a communication center agent's duties with regard to communication center 51. For example, station 79 has implemented therein a personal computer/video display unit (PC/VDU) 82 and an agent's telephone 83. Communication station 81 is likewise equipped with a PC/VDU 84 and an agent's telephone 85. Agent's telephones 83, and 85 are connected to switch 69 via internal wiring 75 as is known in the art. There may be many more agent stations than the two illustrated.

Communication stations 79 and 81 are interconnected via their PC/VDU's to a LAN 77 for the purpose of obtaining and sharing information through the course of normal communication-center operation. A customer information system (CIS) server 87 is connected to LAN 77 and provides a source of information regarding customers, products, services, and other like information. Processor 71 also is LAN connected.

It will be apparent to one with skill in the art that there will be many more communication stations such as station 79 operating in an actual

purposes regardless of what type of queue the calls are in, FIFO, Priority, Skill-based, Virtual, and so on.

It should be noted here that queues described wherein priority assignment, skill-based routing, agent status routing, and so on are instituted are still technically regarded as FIFO queues only in the sense that a progression is made with regards to calls coming in to queue and calls being answered from queue. The ability to stack queues and have calls advance according to various enterprise rules as disclosed with reference to co-pending application 09/024,825 would, of course, require more complex algorithms and statistical reporting in order to provide callers with a reasonably accurate EWT. Such capability is not available in prior art systems.

In order to accomplish accurate EWT in an intelligent routing environment such as system 45, the basic formula used for a simple FIFO ACD queue must be expanded, and better statistical handling and reporting must be observed as described above. In a preferred embodiment, statistical analysis and reporting of call behavior is provided via Stat-server shown as part of software 89. Stat-server software can be adapted to monitor and provide statistics regarding queues, switches, agent status, call traffic, and so on. This method is vastly superior over prior art. Statistical compilation capability may also be extended into PSTN 47 via digital network 63 and processor 57. Similarly, EWT and T-server capability may also be extended into PSTN 47 via the same conventions. In this way EWT may be provided at network level queues associated with SCP 53. More detail regarding expanded mathematical formulas and application thereof to various queue situations for practicing EWT is provided below. Also hierarchical systems can de assembled, where several call centers are connected to a network, and controlled by a common SCP.

are averaging 160 seconds per call, and certified traders who speak Spanish are averaging 170 seconds per call.

The above CDT figures are real-time numbers based on statistical reporting provided by Stat-server software as part of software 89 of Fig. 2. In a preferred embodiment, CDT averages take into account the rate of abandoned calls occurring within the queue and the amounts of time an agent may spend taking calls from another queue if there is more than one queue. These factors are randomly occurring events and are therefore impossible to account for when using the basic formula as described above.

An information table 105 lists some additional factors which can effect an EWT determination for an incoming call. These are abandoned calls (described above), bumped calls (priority queue), re-directed calls (error routed or transferred), use of multiple queues, and use of virtual queues.

Calls are stacked in queue according to priority and skill requirement of a caller. For example, in the column under title-block "Highest", the calls having the highest priority are listed according to skill requirement. To the right, columns labeled 2-7 and "lowest" reflect incremental lower levels of call priority with actual calls waiting listed according to skill requirement. For example, the lowest priority column has 5 calls listed and waiting for English speaking non-certified agents. There are 7 calls ahead of the 5 lowest priority calls. These are 3 calls in the fourth priority column, 2 calls in the third priority column, and 2 calls in the highest priority column. In this embodiment, calls having a same priority assignment in queue are answered according to FIFO rules, however a new call assigned a higher priority would be placed ahead of any lower priority calls in queue and behind any higher priority calls.

In this notation, a fraction a_i represents the fraction of time an agent spends answering calls from a single queue *i*. These fractions (may vary with each agent) must be summed up over all of the agents answering calls from the queue. This result represents the *effective* number of agents for the calculations used as m in equation I.

When taking into account an abandoned call factor which is a random factor of EWT itself, the above notation is multiplied by the percentage of calls that are not expected to be abandoned as follows:

$$\text{III. EWT} = ((1 - r_a)(N \times T_h)) / \sum_{i=1}^{i=m} a_i$$

In this equation, r_a is the rate of abandoned calls, so $(1 - r_a.)$ is the rate of not-abandoned calls. This rate is computed as a dynamically self adjusting factor which takes into account historical information on abandoned calls obtained from Stat-server statistics. It will be apparent to one skilled in the art of statistical calculation that the accuracy of this statistic will improve as more information on call behavior becomes available. This is but one example of how separate gathering of information by Stat-server software of software 89 of Fig. 2 is superior to prior art methods.

With the power of compiling statistical information concerning call behavior such as CDT, rate of abandoned calls, rate of calls bumped, swapped or redirected calls, and so on, a certain confidence level regarding the accuracy of these figures may be developed through further calculation. These calculations are, in a preferred embodiment, performed via EWT software in conjunction with Stat-server software of software 89 of Fig. 2.

her call is next to be answered. More description regarding a virtual queue is provided with regard to co-pending patent application 09/024,825. The methods of the present invention may be employed with virtual queues with a high degree of success as long as the designer of the routing strategy provides meaningful objects from which statistics will be measured. Because of the integrative nature of components of software 89 of Fig. 2 namely, EWT, Stat-server, and T-server implementations, new routing strategies employing virtual queues may selectively obtain historical statistics related to the behavior of specific types of calls that may be generic to the new strategy.

Predicting Optimum Out-Bound Call Generation

As described in the background section, treating calls resulting from an out-bound campaign should be dynamically different from treating normal inbound calls because of a basic fact that in the case of outbound calls, customers will not likely accept significant wait periods in queue. The inventor provides a method for determining optimum frequency for outbound call generation that meets the following two conditions. One is that agents should not be under-utilized. The other is that the rate of call abandonment in queue should be less than an acceptable maximum.

Fig. 4 is a block diagram illustrating out-bound call processing as known to the inventor. A communication center 107 is adapted for initiating out-bound call campaigns and routing connected calls as incoming events to available agents. Center 107 may be assumed to be analogous to center 51 of Fig. 2 above in terms of equipment and functionality. However, the description of this embodiment will focus on the aspect of an out-bound customer-care campaign launched from within the center.

In a preferred embodiment, queue 117 is a first-in-first-out (FIFO) queue. However, in alternative embodiments more advanced queuing techniques may be employed. The inventor chooses a FIFO queue in this embodiment because of a presumption that customers called during an outbound campaign will not accept any significant wait periods in queue 117. Only calls answered by a live person are queued in this example. Moreover, IVR functionality may be assumed to be present in this example though none is illustrated.

The outbound system described above is enhanced by availability of statistical and actual communication-center information provided by virtue of *state* monitoring and compilation of data performed by an instance of STAT server previously described in disclosure pertaining to cross referenced S/N 09/209,306. Therefore, the present invention seeks to incorporate a number of differing variables and constraints, determined through historical and real-time analysis, into algorithmic function in order to obtain an optimum determination of call-generation frequency for an outbound campaign. These variables are described and illustrated below.

Fig. 5 is a block diagram illustrating variables considered in determining call-generation frequency according to an embodiment of the present invention. This example diagram emulates the example of Fig. 4 above with the exception of the addition of a Stat-server 125 and related data flows. Therefore, elements that are present in this example and also in Fig. 4 retain the same element numbers given in Fig. 4.

In the outbound system described in Fig. 4, and also represented herein, generator 113 provides telephone numbers to preprocessor 109, which passes live connections to queue 117 subscribed to by agents 111 who are, in this example, dedicated to the outbound campaign. A directional CIP

defined as the number of calls waiting in queue at any given time. The other value illustrated is actual time in queue (ATQ) per call. ATQ refers to calls that leave queue 117 without being served by an agent.

A directional arrow illustrated as leading from preprocessor 109 and progressing toward server 125 represents data reported to server 125 by processor 109. One illustrated value is actual time each call is in preprocessor 109 (ATP). Another is data concerning the capacity of elements (COE) in preprocessor 109. The inventor intends that this example represent just some of the information categories that may be compiled by Stat-server 125 during monitoring of components 109-113. However, there are many other data categories and sub-categories that may be considered by the system of the present invention. Some additional categories that are not illustrated herein may include, but are not limited to, total number of calls in preprocessor, total rate of calls abandoned in preprocessor, rate of calls determined busy in preprocessor, and so on. The invention seeks to utilize key data values in order to develop a flexible method for determining an optimum frequency for call generation as is illustrated further below.

There are two methods that may be used to predict an optimum call-generation frequency for an outbound campaign given a set of variables and conditions. One is by using an analytical approach. The other method is practiced through simulation techniques (object modeling). Both methods are described below in this specification.

__Analytical Approach__

The main goal of applying an analitical approach for predicting call-generation frequencies is to estimate a dialing rate using rather simple formulas from classic queuing theory. It is known from queuing theory that

$$p_k = \frac{a^k/k!}{\sum_{i=0}^{m} a^i/i!}, \; k = \overline{0,m}, \qquad (1)$$

over-dial (abandon rate) that in not greater than Ar_{max}, with a utilization rate no less than ρ_{min}. If Ar_{max} is decreased, the value of ρ_{min} should be also decreased. There is a functional one-to-one relationship between the parameters of the above-mentioned values. Therefore, if one of the values is held fixed, then the other value may be determined through calculation. Thus, the appropriate outbound dialing frequency (call generation frequency) relates (after calculation) to the value pair ($Ar_{max.}$ ρ_{min}).

Referring back to Fig. 4, in order to determine the appropriate rate for call generation for generator 113, it is important to consider a mathematical model of the entire outbound dialing system. It is preferred, as previously described, that calls having progressed into queue 117 are those calls that have been answered by a live person and should not remain in queue 117 for any length of time. In actual practice, time in queue must be only a second or so. This time span is small enough to allow a mathematical representation of the system of Fig. 4 without a queue. For purpose of discussion then, all calls progressing to queue 117 are abandoned if all agents are busy with calls. It is of course assumed in this application that the outbound call flow exhibits stochastic characteristics.

It is important to note here that a solution to determining a proper call generation frequency for generator 13 is a predictive one. In observance of Erlang's formula, it is noted that p_k is the probability that k servers (agents) are busy. In application of the basic formula (1) described above, consider that $a = \lambda / \Phi$ where λ is the average rate of the input call flow of the system and Φ is the average rate of the output flow. The present invention applies a solution under such condition as $\lambda = (n\ \Phi) < 1$. Without the just-mentioned condition or constraint, input traffic would be more than output 1. In that case, call abandon rate would increase

value will reveal the correct frequency for call generation observing the constraint Ar_{max}. Steps to solving the problem include:

1. Estimating values p, λ_i, and φ. These values can be calculated using current data. The value of $\varphi = 1/t$, where t is the average duration of an answered call.

2. Obtaining the value a_m as a solution from equation (2) above with respect to a.

3. Obtaining the value of λ_m from equation (3) above wherein $\lambda_m = (a_m\Phi - \lambda_i)/p$.

One with skill in the art of mathematical expression and functional notation will recognize the functionality provided by the above notations.

Optimizing Agent Utilization

Agent utilization or busy factor will be expressed as B. This value represents an average number of busy (engaged in active calls) agents in equilibrium using the model represented by Figs. 4 and 5. Using Erlang's formula, B is expressed as follows:

$$B = B(a) = \sum_{k=1}^{m} k \frac{a^k/k!}{\sum_{i=0}^{m} a^i/i!} = \frac{aR(m-1,a)}{R(m,a)}, \qquad (4)$$

where $R(m,a) = \sum_{k=0}^{m} a^k/k!$.

It is exemplified by the above formula that B is a function of a. Here parameter a has the same mean as above ($a = \lambda/\Phi$). In order to solve the

Where $F(x,a)$ is a probability distribution function for normal random variables with a mathematical expectation a and a variance a. It is well known that $F(x,a)$ may be expressed as follows:

$$F(x,a) = \frac{1}{\sqrt{2\pi a}} \int_{-\infty}^{x} e^{-(z-a)^2/(2a)} dz.$$

The formula (7) illustrated further above can be easily derived from the following facts. The Erlang's formula gives the probability distribution for Poisson random variable ξ under condition $\xi \leq m$. If m is large enough, the value of a_m for a reasonable value of Ar_{max} or ρ_{min} is also large. Hence we can use the normal approximation for the Poisson distribution, which is well known in the art.

Simulation Modeling

In another aspect of the present invention, a simulation (object modeling) method is used to solve the call generation frequency problem while continuing to observe the stated optimization goals, which were described above.

In one aspect, a model represents a simplified scope of a whole outbound call-generation and service system. The same model equates to a strict and formal representation of the system. This fact allows formal treatments including logical reasoning about its properties and therefore about properties of the whole system. A simulation model of an outbound

- m - the total number of agents;
- ρ_{min} - the minimal agents' utilization rate;
- Ar_{max} - the maximal abandon rate.
- P – agent utilization rate.
- Ar – abandon rate.

The following assumptions are attributed to distribution function of model 127:

1) Agents' service times have exponential distribution with parameters τ_1 and τ_2:

$$f_1(t) = \begin{cases} \frac{1}{\tau_1} e^{-\frac{t}{\tau_1}}, & \text{if } t \geq 0; \\ 0, & \text{otherwise.} \end{cases} \quad \text{and} \quad f_2(t) = \begin{cases} \frac{1}{\tau_2} e^{-\frac{t}{\tau_2}}, & \text{if } t \geq 0; \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

2) A call abandonment process has an exponential distribution with parameter μ:

3) The preprocessor delays calls relating to a deterministic delay process having a delay parameter τ_3.

In defining a precise statement of the problem to be solved, consider that t_0 is an initial moment of time. At this moment there exists a current number of busy agents K_0, a current queue length Q_0, and a current number of calls still in preprocess mode P_0. Also consider that for each call still at preprocessor (109), there is a rest time defined as the period of time that call remains in preprocessor (109) or tr_i where $0 \leq i \leq P_0$. The solution is to

state implies that there may be no way to fully satisfy all of the requirements in a stochastic call flow. Therefore, a more general approach must be employed with respect to simulation.

$$\Phi\ (Ar,\rho) \tag{10}$$

The above function is optimized by observing two states:

1) Abandon rate (Ar) should be maintained at admissible level while maximizing agent utilization rate (ρ)is as much as possible.

2) Agent utilization rate (ρ) is maintained at an acceptable level while abandonment ratio is held as small as possible.

In creation of a simulation with respect to agent parameters, a mathematical starting point is observed for both queue length and for the number of busy agents. The first formula is closely related to a call abandon rate and the second formula is related to agent utilization. The process defining the length of a queue has the following form:

$$Q_t = Q_0 + \int_0^t I\{K_s = m\}dN_s - D_t - S_t, \tag{11}$$

Where $Q0$ is a queue length at an initial moment of time;
Control process N effects a jump at moment of entry of a next call entering queue from a preprocessor. xxx

An additional formula Dt =tR0_Qsds represents an abandonment of a client while in queue without benefit of agent service.

A formula is derived for expressing abandon rate Ar. Abandon rate is defined as a ratio of abandoned calls to all calls left in queue during some time interval and before agent processing. Clearly $0 \leq Ar \leq 1$. $Q_1 - Q_0 = \Delta Q$ or the noted change in a queue during time interval ΔT, which = $t_1 - t_0$. ΔQ_D is a number of calls left in a queue due to abandonment during ΔT. Then formula for the average abandon rate during the interval ΔT can be expressed as follows:

$$\overline{Ar} = \frac{\Delta Q_D}{\Delta Q} \qquad (15)$$

To calculate the average agent utilization rate ρ at the interval ΔT we will use the formula:

$$\overline{Ar} = \frac{\Delta Q_D}{\Delta Q} \qquad (15)$$

Where K_i is the average number of busy agents at the interval $\Delta t_i = t_i+1 - t_i$ and n is a number of intervals Δt_i within the interval ΔT.

The actual simulation programming language is not detailed herein as one with skill in the art will recognize and appreciate the process of simulation based on queuing theory. In a preferred embodiment C++ programming language is used.

The novel portion of the programming method is that it takes into account the conditions that were mentioned further above in this specification by simulating each of the processes and allowing a prediction of

What is claimed is:

1. A system for balancing outbound dialing rate with agent utilization in a telephony call center, minimizing wait time for answered outdialed calls, comprising:

a call number generating module for generating numbers to be automatically dialed;

a dialing unit for dialing numbers generated;

a queue for queing answered calls; and

a stat module for monitoring performance and generating a call generation rate;

characterized in that the system uses an analytical method for determining the call generation rate, wherein the method uses one or more of distribution function of system processes, numerical methods, solving of non-linear equations, or probability techniques.

2. An analytical method for minimizing time in queue for parties answering outdialed calls, comprising steps of:

(a) monitoring elements of system behavior using a stat module connected to a call generating unit;

(b) calculating a call generation rate by applying one or more of distribution function of system processes, numerical methods, solving of non-linear equations, or probability techniques with statistics developed by the stat module; and

(c) applying the call generation rate to pace the call generation unit.

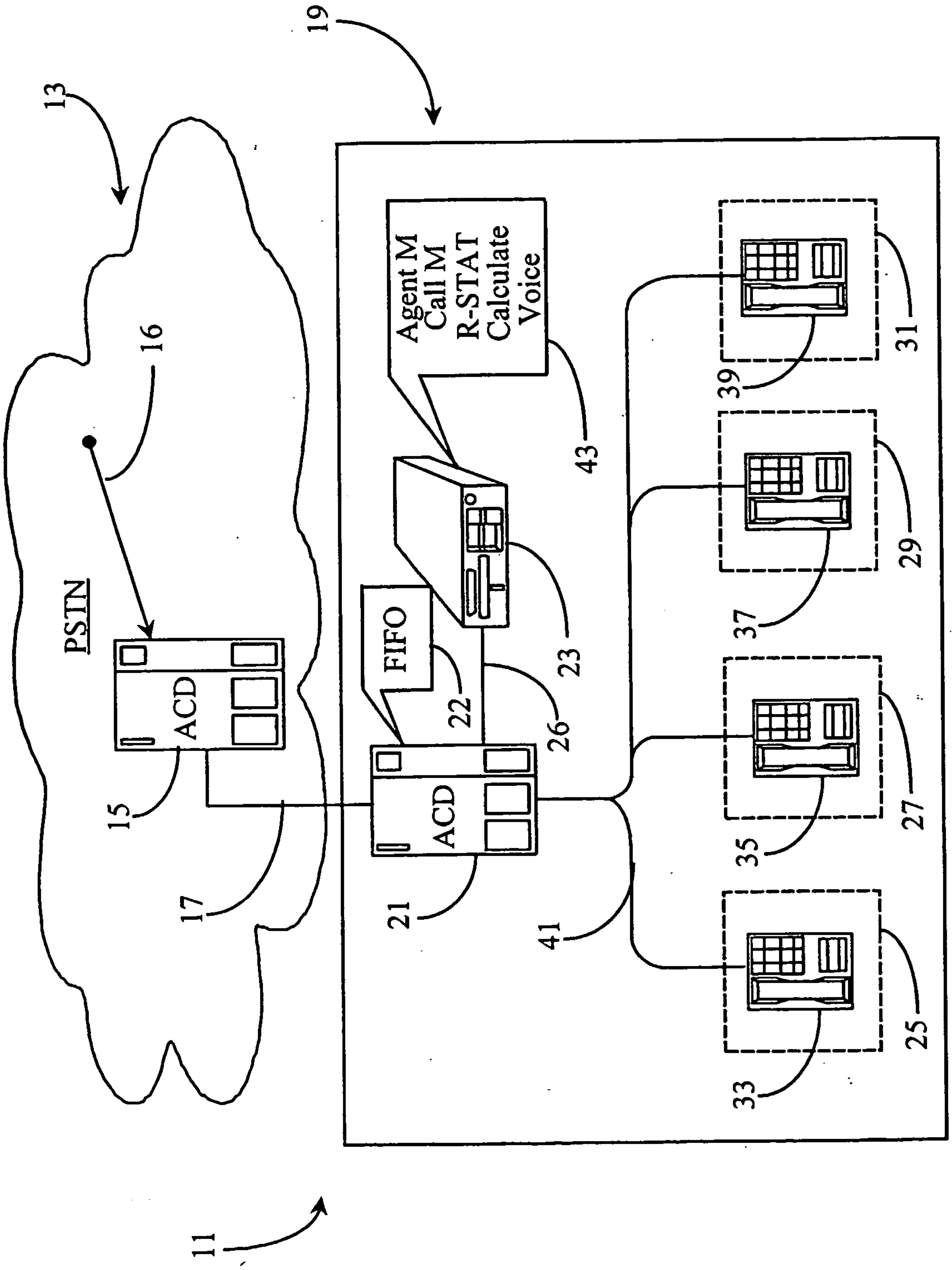


Fig. 1 (prior art)

91 103 101

Agent Skill	# of Agents	AVG Call	Highest	2	3	4	5	6	7	Lowest
English	3	120/Sec	2	0	2	3	0	0	0	5
Spanish	2	100/Sec	1	3	1	0	0	0	0	7
Certified Trader/Eng.	3	160/Sec	1	0	5	0	0	0	2	0
Certified Trader/Sp.	1	170/Sec	1	0	0	2	0	0	0	0

93 95 97 99 105

Abandoned Calls
Bumped Calls
Redirected Calls
Multiple Queues
Virtual Queues

Skill-Based Priority

Fig. 3

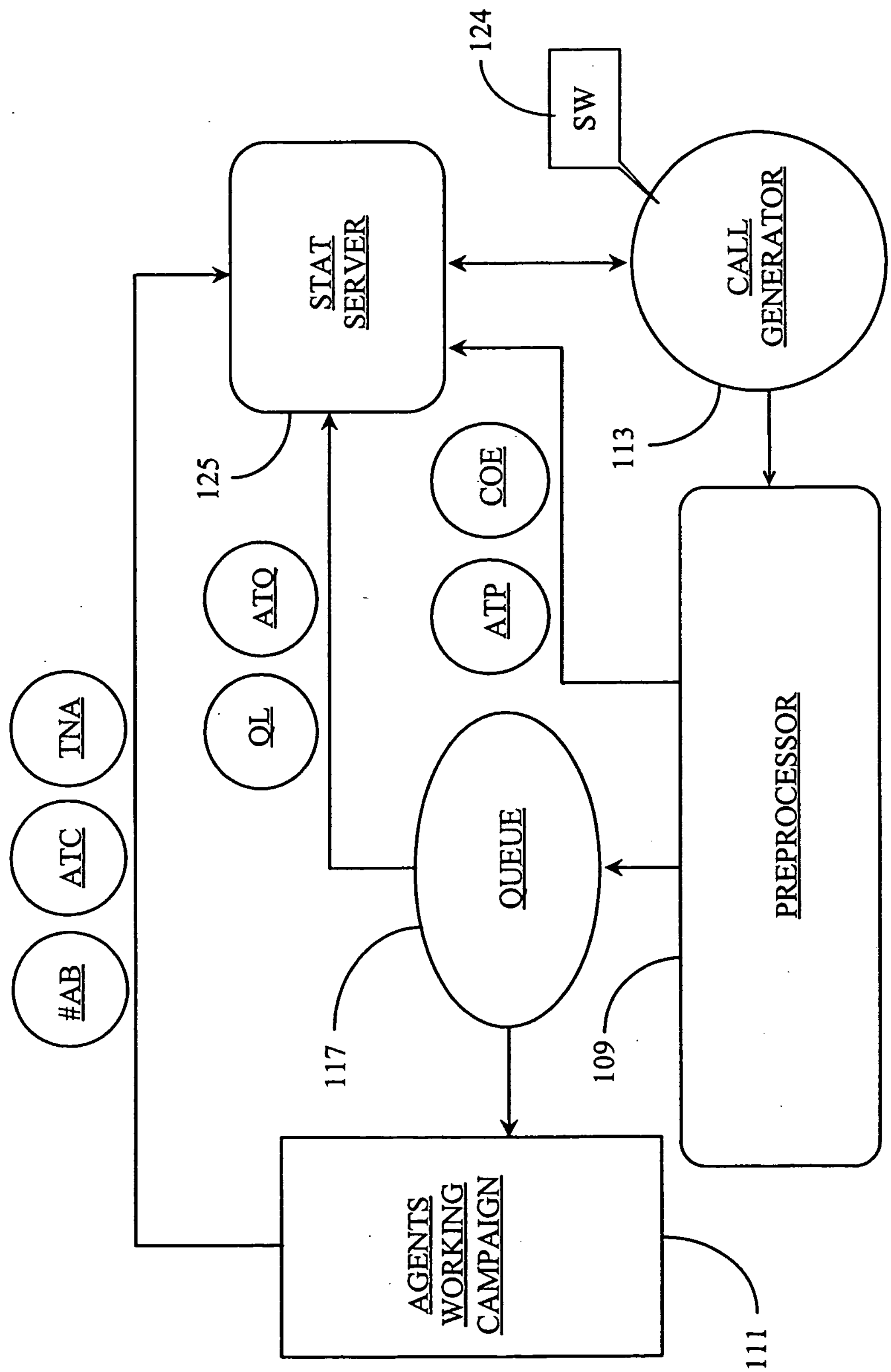

STAT SERVER
CALL GENERATOR
SW
124
125
113
COE
ATO
ATP
QL
TNA
ATC
#AB
QUEUE
PREPROCESSOR
117
109
AGENTS WORKING CAMPAIGN
111

Fig. 5

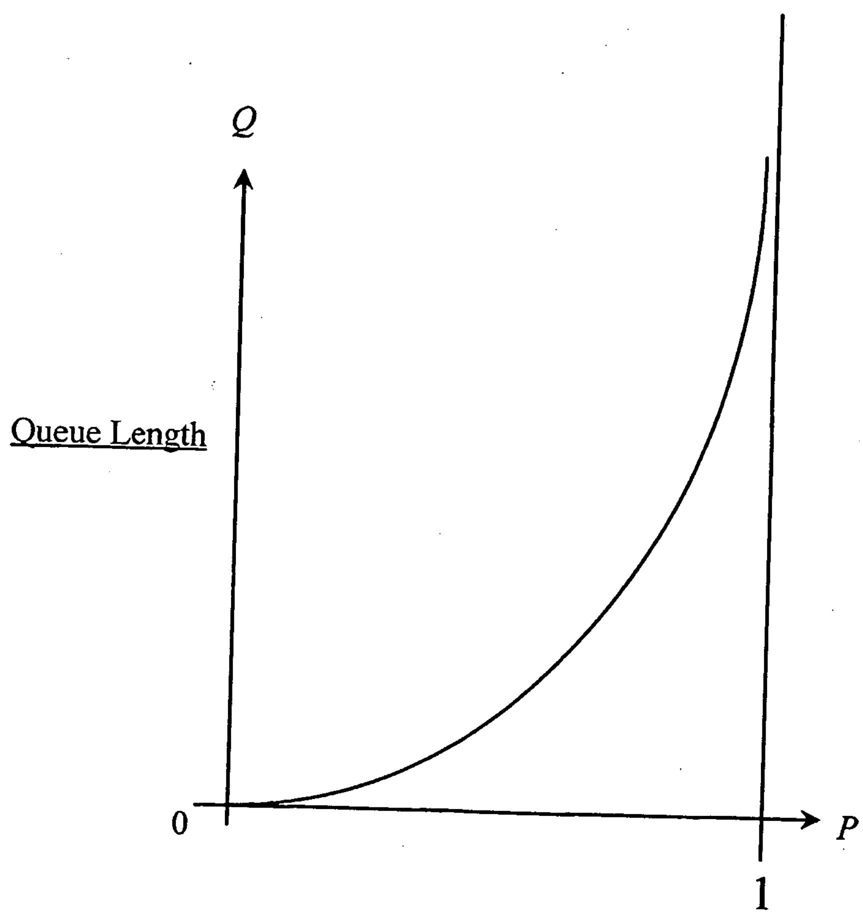


Fig. 7.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/07457

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :H04M 3/00
US CL :379/265
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 379/265, 34, 215, 242, 266, 268-269, 272, 308-309, 904

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EAST
search terms: ACD, outbound, queue$, wait$, dial$, rate$

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4,881,261 A (OLIPHANT et al) 14 November 1989, see abstract, summary, col. 7, lines 30-51	1-4
A	US 5,247,569 A (CAVE) 21 September 1993, see abstract, summary, col. 1, lines 33-65	1-4
A	US 5,467,391 A (DONAGHUE, Jr. et al) 14 November 1995, see abstract, col. 1, line 61 - col. 2, line 5, col. 5, lines 37-53.	1-4

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 MAY 2001	30 AUG 2001

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	WELLINGTON CHIN
Facsimile No. (703) 305-3230	Telephone No. (703) 305-4366